AFRL-OSR-VA-TR-2013-0087

Specification-based Error Recovery: Theory, Algorithms, and Usability

Sarfraz Khurshid
University of Texas at Austin

FEBRUARY 2013
Final Report

AIR FORCE RESEARCH LABORATORY
AF OFFICE OF SCIENTIFIC RESEARCH (AFOSR)/RTC
ARLINGTON, VIRGINIA 22203
AIR FORCE MATERIEL COMMAND

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
02-03-2013	Final Report	May 1, 2009 - June 30, 2012

4. TITLE AND SUBTITLE

Specification-based Error Recovery: Theory, Algorithms, and Usability

5a. CONTRACT NUMBER

FA9550-09-1-0351

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Sarfraz Khurshid
Associate Professor, Electrical and Computer Engineering

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Texas at Austin
1 University Station
Austin, TX 78712

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Office of Scientific Research
875 N Randolph St
Arlington, VA 22203
Dr. Robert Bonneau/RTC

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

AFRL-OSR-VA-TR-2013-0087

12. DISTRIBUTION/AVAILABILITY STATEMENT

Distribution A: Approved for public release.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This project laid the foundation for a novel methodology for correcting erroneous program executions using specifications at run-time. The basis of the methodology is a view of the specification as a non-deterministic implementation, which may permit a high degree of non-determinism. The key insight is to use likely correct actions by an otherwise erroneous execution to prune the non-determinism in the specification, thereby transmuting the specification to an implementation at run-time and reducing the performance overhead. A suite of techniques and tools were designed, developed, optimized and rigorously evaluated in this project. It leveraged the Alloy specification language and its SAT-based tool-set as an enabling technology for specification-based analysis. The ideas, techniques, tools, and evaluation results from this project contributed in part to archival publications, Masters theses, and PhD dissertations.

15. SUBJECT TERMS

Data structure repair, error recovery, specification-based analysis, program repair

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Sarfraz Khurshid
Unclassified	Unclassified	Unclassified	UU	9	19b. TELEPHONE NUMBER *(Include area code)* (512) 471-8244

Reset

Specification-based Error Recovery: Theory, Algorithms, and Usability

FA9550-09-1-0351

Final Report (May 1, 2009 to June 30, 2012)

Principal Investigator: Sarfraz Khurshid
PI Email: khurshid@ece.utexas.edu
PI Tel: (512) 471-8244
Program Manager: Dr. Robert Bonneau

1 Summary

This project laid the foundation for a novel methodology for correcting erroneous program executions using specifications at run-time. The basis of the methodology is a view of the specification as a non-deterministic implementation, which may permit a high degree of non-determinism. The key insight is to use likely correct actions by an otherwise erroneous execution to prune the non-determinism in the specification, thereby transmuting the specification to an implementation at run-time and reducing the performance overhead. A suite of techniques and tools were designed, developed, optimized and rigorously evaluated in this project. It leveraged the Alloy specification language and its SAT-based tool-set as an enabling technology for specification-based analysis. The ideas, techniques, tools, and evaluation results from this project contributed in part to 44 archival publications, 4 completed Masters theses, and 3 completed PhD dissertations. This project funded in part 8 graduate students, including 3 female students.

2 Annual summaries

2.1 Reporting period: 05/01/2009 – 04/30/2010

During the first year of the project, the following research contributions were made:

- Contract-based data structure repair – Introduced the idea of using rich behavioral contract specifications including invariants, pre- and post-conditions as the basis of systematic data structure repair.
- Repair algorithms – Developed four algorithms that embody the idea. The algorithms leverage MIT's Alloy tool-set to provide systematic repair, and employ heuristics to optimize performance.
- Similarity metric – Used a distance metric for graph similarity to compute the effect of repair on an erroneous program state and to evaluate different algorithms for effectiveness.
- Evaluation – Conducted an experimental evaluation of the feasibility of contract-based repair and demonstrated the promise it holds.

A basic technique embodying these ideas and and experimental evaluation were presented at the 24th European Conference on Object-Oriented Programming (ECOOP) in June 2010; a pre-print version of the paper is submitted along with this report.

2.2 Reporting period: 05/01/2010 – 04/30/2011

During the second year of the project, our primary research contribution was on program repair using data structure repair. A key element of the "Usability" thrust of our project is to design a repair feedback mechanism to help users debug their code or specifications. We developed a novel mechanism for translating repair actions performed on an erroneous program state into code that abstracts those actions using assignment statements that may replace existing program statements or be added as new statements. These statements serve as debugging suggestions, which the user can choose to apply or ignore. Details of this approach and an experimental evaluation were presented at the IEEE 4th International Conference on Software Testing, Verification and Validation (ICST) in March 2011; a pre-print version of the paper is submitted along with this report.

2.3 Reporting period: 05/01/2011 – 04/30/2012

During the third year of the project (May 1, 2011 to April 30, 2012), our primary research contribution was to develop a new technique to enhance our core approach for data structure repair to scale better. Our insight into scalability is two-fold: (1) the dynamic program trace of field writes and reads provides useful guidance to repair incorrect state mutations by a

faulty program; and (2) unsatisfiable cores generated by SAT can capture the history of previous runs, which can be used in an efficient iterative approach on successive problems with increasing state spaces. Details of this technique and an experimental evaluation were presented at the 18th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS) in March 2012; a pre-print version of the paper is submitted along with this report.

Additionally, we utilized unsatisfiable cores in another novel technique, which was for fault localization – the problem of locating faults in the source-code of a buggy program. Specifically, we developed a specification-based technique that utilized correct and erroneous executions of the buggy program to more accurately locate faults. Our insight is that unsatisfiability analysis of violated specifications, enabled by SAT technology, can help (1) compute unsatisfiable cores that contain likely faulty statements; and (2) generate tests that help spectra-based localization. Details of this technique and an experimental evaluation were presented at the 27th IEEE/ACM International Conference on Automated Software Engineering (ASE) in September 2012; a pre-print version of the paper is submitted along with this report.

2.4 Reporting period: 05/01/2012 – 06/30/2012

During the final two months of the project (May 1, 2012 to June 30, 2012), we focused on enhancing the ideas, analyses, and implementations we developed in this project to integrate them as parts of doctoral dissertations – we expect three future doctoral dissertations to use the work done in this project at their foundation.

3 Archival publications

3.1 Published after the end of the funding period

1. R. N. Zaeem and S. Khurshid. Test Input Generation Using Dynamic Programming. In *Proc. ACM SIGSOFT 20th International Symposium on the Foundations of Software Engineering (FSE)*, 11 pages, Research Triangle Park, NC, Nov. 2012

2. L. Zhang, M. Kim, and S. Khurshid. FaultTracer: A Change Impact and Regression Fault Analysis Tool for Evolving Java Programs. In *Proc. ACM SIGSOFT 20th International Symposium on the Foundations of Software Engineering (FSE)*, 4 pages, Research Triangle Park, NC, Nov. 2012. Research tool demonstration paper

3. C. H. P. Kim, S. Khurshid, D. Batory. Shared Execution for Efficiently Testing Product Lines. In *Proc. IEEE International Symposium on Software Reliability Engineering (ISSRE)*, Dallas, TX, Nov. 2012

4. S. Ganov, S. Khurshid, and D. E. Perry. Annotation-aided Automated Incremental Analysis for Alloy via Domain Specific Solvers. In *Proc. 14th International Conference on Formal Engineering Methods (ICFEM)*, pages 414–429, Kyoto, Japan, November 2012

5. S. Roychowdhury and S. Khurshid: Localization of faults in software programs using Bernoulli divergences. In *Proc. International Symposium on Information Theory and its Applications (ISITA)*, pages 586–590, Honolulu, HI, Oct. 2012

6. J. H. Siddiqui and S. Khurshid. Scaling symbolic execution using ranged analysis. In *Proc. ACM International Conference on Object Oriented Programming Systems Languages and Applications (OOPSLA)*, pages 523–536, Tuscon, AZ, Oct. 2012

7. S. Roychowdhury. Ensemble of feature selectors for software fault localization. In *Proc. IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 1351–1356, Seoul, Korea, Oct. 2012

8. S. Roychowdhury and S. Khurshid. A family of generalized entropies and its application to software fault localization. In *Proc. IEEE Conference on Intelligent Systems (IS)*, pages 368–373, Sofia, Bulgaria, Sep. 2012

9. D. Gopinath, R. N. Zaeem, and S. Khurshid. Improving the Effectiveness of Spectra-based Fault Localization using Specifications. In *Proc. 27th IEEE/ACM International Conference on Automated Software Engineering (ASE)*, pages 40–49, Essen, Germany, Sep. 2012

10. G. Yang, S. Khurshid, and M. Kim. Specification-based test repair using a lightweight formal method. In *Proc. 18th International Symposium on Formal Methods (FM)*, pages 455–470, Paris, France, Aug. 2012

11. L. Zhang, D. Marinov, L. Zhang, and S. Khurshid. Regression mutation testing. In *Proc. International Symposium on Software Testing and Analysis (ISSTA)*, pages 331–341, Minneapolis, MN, July 2012

12. G. Yang, C. Pasareanu, and S. Khurshid. Memoized symbolic execution. In *Proc. International Symposium on Software Testing and Analysis (ISSTA)*, pages 144–154, Minneapolis, MN, July 2012

3.2 Published during the funding period

13. M. Z. Malik and S. Khurshid. Dynamic shape analysis using spectral graph properties. In *Proc. IEEE Fifth International Conference on Software Testing, Verification and Validation (ICST)*, pages 211–220, Montreal, Canada, Apr. 2012

14. J. H. Siddiqui, D. Marinov, and S. Khurshid. Lightweight data-flow analysis for execution-driven constraint solving. In *Proc. IEEE Fifth International Conference on Software Testing, Verification and Validation (ICST)*, pages 91–100, Montreal, Canada, Apr. 2012

15. R. N. Zaeem, D. Gopinath, S. Khurshid, and K. S. McKinley. History-aware data structure repair using SAT. In *Proc. 18th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS)*, pages 2–17, Tallinn, Estonia, Mar. 2012

16. J. H. Siddiqui and S. Khurshid. Staged symbolic execution. In *Proc. ACM Symposium on Applied Computing (SAC)*, pages 1339–1346, Riva del Garda, Italy, Mar. 2012

17. L. Zhang, D. Marinov, L. Zhang, and S. Khurshid. An empirical study of JUnit test-suite reduction. In *Proc. 22nd International Symposium on Software Reliability Engineering (ISSRE)*, pages 170–119, Hiroshima, Japan, Nov. 2011

18. S. A. Khalek and S. Khurshid. Efficiently running test suites using abstract undo operations. In *Proc. 22nd International Symposium on Software Reliability Engineering (ISSRE)*, pages 110–119, Hiroshima, Japan, Nov. 2011

19. S. Roychowdhury and S. Khurshid. Software fault localization using feature selection. In *Proc. International Workshop on Machine Learning Technologies in Software Engineering (MALETS)*, pages 11–18, Lawrence, KS, Nov. 2011

20. S. A. Khalek, G. Yang, L. Zhang, D. Marinov, and S. Khurshid. TestEra: A tool for testing Java programs using Alloy specifications. In *Proc. 26th IEEE/ACM International Conference on Automated*

Software Engineering (ASE), pages 608–611, Lawrence, KS, Nov. 2011. Tool Demo Paper

21. S. A. Khalek, V. P. Narayanan, and S. Khurshid. Mixed constraints for test input generation – An initial exploration. In *Proc. 26th IEEE/ACM International Conference on Automated Software Engineering (ASE)*, pages 548–551, Lawrence, KS, Nov. 2011. Short paper

22. S. R. Ganov, S. Khurshid, and D. E. Perry. A case for Alloy annotations for efficient incremental analysis via domain specific solvers. In *Proc. 26th IEEE/ACM International Conference on Automated Software Engineering (ASE)*, pages 464–467, Lawrence, KS, Nov. 2011. Short paper

23. J. H. Siddiqui and S. Khurshid. Symbolic execution of Alloy models. In *Proc. 13th International Conference on Formal Engineering Methods (ICFEM)*, pages 340–355, Durham, UK, Oct. 2011

24. L. Zhang, M. Kim, and S. Khurshid. Localizing failure-inducing program edits based on spectrum information. In *Proc. IEEE 27th International Conference on Software Maintenance (ICSM)*, pages 23–32, Williamsburg, VA, Sept. 2011

25. S. Roychowdhury and S. Khurshid. A novel framework for locating software faults using latent diver- gences. In *Proc. European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD)*, pages 49–64, Athens, Greece, Sept. 2011

26. S. Person, G. Yang, N. Rungta, and S. Khurshid. Directed incremental symbolic execution. In *Proc. 32nd ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI)*, pages 504–515, San Jose, CA, June 2011

27. C. Cadar, P. Godefroid, S. Khurshid, C. S. Pasareanu, K. Sen, N. Tillmann, and W. Visser. Symbolic execution for software testing in practice: Preliminary assessment. In *Proc. 33rd International Conference on Software Engineering (ICSE)*, pages 1066–1071, Waikiki, Honolulu, May 2011

28. Muhammad Zubair Malik. Dynamic Shape Analysis of Program Heap using Graph Spectra (NIER Track). In *Proc. 33rd International Con-*

ference on Software Engineering (ICSE), pages 952–955, Waikiki, Honolulu, May 2011

29. D. Gopinath, M. Z. Malik, and S. Khurshid. Specification-based program repair using SAT. In *Proc. 17th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS)*, pages 173–188, Saarbrucken, Germany, Mar. 2011

30. S. A. Khalek and S. Khurshid. Systematic testing of database engines using a relational constraint solver. In *Proc. IEEE Fourth International Conference on Software Testing, Verification and Validation (ICST)*, pages 50–59, Berlin, Germany, Mar. 2011

31. M. Z. Malik, J. H. Siddiqui, and S. Khurshid. Constraint-based program debugging using data structure repair. In *Proc. IEEE Fourth International Conference on Software Testing, Verification and Validation (ICST)*, pages 190–199, Berlin, Germany, Mar. 2011

32. E. Uzuncaova, S. Khurshid, and D. Batory. Incremental test generation for software product lines. *IEEE Transactions on Software Engineering (TSE)*, 36(3):309–322, 2010

33. D. Shao, D. Gopinath, S. Khurshid, and D. E. Perry. Optimizing incremental scope-bounded checking with data-flow analysis. In *Proc. IEEE 21st International Symposium on Software Reliability Engineering (ISSRE)*, pages 408–417, San Jose, CA, Nov. 2010

34. J. H. Siddiqui and S. Khurshid. ParSym: Parallel symbolic execution. In *Proc. 2nd International Conference on Software Technology and Engineering (ICSTE)*, pages V1-405–V1-409, San Juan, PR, Oct. 2010

35. S. A. Khalek and S. Khurshid. Automated SQL query generation for systematic testing of database engines. In *Proc. 25th IEEE/ACM International Conference on Automated Software Engineering (ASE)*, pages 329–332, Antwerp, Belgium, Sept. 2010

36. R. N. Zaeem and S. Khurshid. Contract-based data structure repair using Alloy. In Proc. *24th European Conference on Object-Oriented Programming (ECOOP)*, pages 577–598, Maribor, Slovenia, June 2010

37. M. Gligoric, T. Gvero, V. Jagannath, S. Khurshid, V. Kuncak, and D. Marinov. Test generation through programming in UDITA. In Proc. *32nd ACM/IEEE International Conference on Software Engineering*

(ICSE), pages 225–234, Cape Town, South Africa, May 2010. **ACM SIGSOFT Distinguished Paper Award**

38. R. N. Zaeem and S. Khurshid. Introducing specification-based data structure repair using Alloy. In *Proc. Second International Conference on Abstract State Machines, Alloy, B and Z (ABZ)*, pages 398–399, Orford, Canada, Feb. 2010. Abstract paper

39. D. Shao, D. Gopinath, S. Khurshid, and D. E. Perry. A case for using data-flow analysis to optimize incremental scope-bounded checking. In *Proc. Second International Conference on Abstract State Machines, Alloy, B and Z (ABZ)*, pages 392–393, Orford, Canada, Feb. 2010. Abstract paper

40. J. H. Siddiqui and S. Khurshid. An empirical study of structural constraint solving techniques. In *Proc. 11th International Conference on Formal Engineering Methods (ICFEM)*, pages 88–106, Rio de Janeiro, Brazil, 2009

41. S. R. Ganov, C. Killmar, S. Khurshid, and D. E. Perry. Event listener analysis and symbolic execution for testing GUI applications. In *Proc. 11th International Conference on Formal Engineering Methods (ICFEM)*, pages 69–87, Rio de Janeiro, Brazil, 2009

42. M. Z. Malik, K. Ghori, B. Elkarablieh, and S. Khurshid. Automated debugging using data structure repair. In *Proc. 24th IEEE/ACM International Conference on Automated Software Engineering (ASE)*, pages 620–624, Auckland, New Zealand, Nov. 2009. Short paper.

43. J. H. Siddiqui, D. Marinov, and S. Khurshid. Optimizing a structural constraint solver for efficient software checking. In *Proc. 24th IEEE/ACM International Conference on Automated Software Engineering (ASE)*, pages 615–619, Auckland, New Zealand, Nov. 2009. Short paper.

44. D. Shao, S. Khurshid, and D. E. Perry. An incremental approach to scope-bounded checking using a lightweight formal method. In *Proc. 16th International Symposium on Formal Methods (FM)*, pages 757–772, Eindhoven, the Netherlands, Nov. 2009

4 Masters theses – finished during the funding period

1. Sowmiya Chocka Narayanan. *Clustered Test Execution using Java Pathfinder*. Masters thesis, Department of Electrical and Computer Engineering, University of Texas at Austin, May 2010

2. Divya Gopinath. *Scaling scope bounded checking using incremental approaches*. Masters thesis, Department of Electrical and Computer Engineering, University of Texas at Austin, May 2010

3. Razieh Nokhbeh Zaeem. *Contract-based data structure repair using Alloy*. Masters thesis, Department of Electrical and Computer Engineering, University of Texas at Austin, May 2010

4. Vidya Narayanan. *Milao: A novel framework for mixed imperative and declarative formulation and solving of structural constraints*. Masters thesis, Department of Electrical and Computer Engineering, University of Texas at Austin, Dec 2009

5 PhD theses – finished during the funding period

1. J. H. Siddiqui. *Improving Systematic Constraint-driven Analysis using Incremental and Parallel Techniques*. PhD thesis, Department of Electrical and Computer Engineering, University of Texas at Austin, May 2012.

2. S. A. Khalek. *Systematic testing using test summaries: Effective and efficient testing of relational applications*. PhD thesis, Department of Electrical and Computer Engineering, University of Texas at Austin, Dec. 2011.

3. D. Shao. *Application of local semantic analysis in fault prediction and detection*. PhD thesis, Department of Electrical and Computer Engineering, University of Texas at Austin, May 2010.